8.0. DATA ANALYSIS METHODS

Suggested methods for analysis of monitoring data collected from the Estero Bay Watershed are described below. Methods are identified for three types of analyses:

- assessment of HSPF model fit (calibration and validation);
- land use/management practices-loading relationships; and
- rainfall-loading relationships.

Additional analyses applicable to collected monitoring data that are also presented include:

- analyses of spatial and temporal relationships; and
- relationships between stressor and response variables.

8.1. Use of Monitoring Data to Assess Model Fit: Calibration and Validation

One important use of the results of the previously presented monitoring program will be to calibrate and validate the HSPF model of the Estero Bay Watershed. Important response variables that will be estimated by the HSPF model, such as freshwater discharge and pollutant load concentrations, will be measured concurrently with the dynamic explanatory variables (e.g., rainfall, insolation, and humidity). Post HSPF processing analyses will then be required to evaluate the fit of the HSPF model predictions of discharge and pollutant load concentrations to the measured data. The following methods identify one approach in which monitoring program results can be used for calibration and validation.

8.1.1. HSPF Model Calibration

The monitoring data would be used to calibrate the HSPF model in two steps. The first step of the calibration process would be to use a subjective approach of plotting the predicted and measured response data with the explanatory variables to assess the fit of the model. This would provide an efficient use of the professional judgement of the HSPF modeling team and the complex set of information available to them. The second step would be to estimate an objective measure of the fit of the model. This would provide a straightforward and quantitative metric of the model which could be readily compared to other HSPF model runs. These two steps would be repeated in sequence until the model calibration is completed.

The first step in using the monitoring data to assess the fit of the model would be to use subjective graphical methods to provide an assessment of model fit and the relationships between explanatory and response variables. It is important to note that time series plots overlaying parallel trends of predicted responses versus time and measured responses versus time (e.g., freshwater discharge), although potentially useful, could likely obscure a lack of fit of the HSPF model. This could occur when the visual patterns of the similar response of predicted and observed values to episodic storm events would obscure smaller yet important biases in the predictions. In order to provide the best subjective tool for assessment of the fit of the model, the predicted values would be plotted against the observed values and overlaid with a 1:1 reference line. Deviations from the 1:1 line would then be readily visible. If the deviations from the 1:1 line are deemed to be important, then the lack of fit of the model could be further isolated by plotting categorical explanatory variables such as season or month as different symbols on the plots. This would aid in determining if the lack of fit is associated with the state of a categorical variable. If the lack of fit could still not be isolated or if there would be a non-linear relationship between the predicted and observed response values, then the residuals of the predicted and measured response variables could be plotted against selected continuous explanatory variables. One would then search for relationships between the response variable residuals and explanatory variables such as rainfall or insolation. The existence of such a relationship would indicate a lack of fit of the model related to that explanatory variable.

In addition to the more subjective graphical assessments of the model fit to the measured data from the monitoring program, it would be useful to develop an objective and quantitative metric of the fit of the HSPF model. Such a measure could be easily tabulated for comparison of model runs and could be presented in project reports in a format which could be easily interpreted. All of the data needed to compute such a statistic would be provided by the predicted response values from the model and the measured response values from the monitoring program. The total variation among the measured values for each of the response variables (e.g., freshwater discharge, TSS concentration, TN concentration) could be expressed as:

$$\sum_{i=1}^{I} \left(R_i - \bar{R} \right)^2$$

where I = the number of measured response samples,

R_i = the measured response value from the ith sample of the monitoring program, and

$\bar{R}_i$ = the mean of the measured response samples.

The total variation among the measured response data could then be split into two parts.

- The total variation among the model predictions would be computed as:

$$\sum_{i=1}^{I} \left(\hat{R}_i - \bar{R} \right)^2$$

where $\hat{R}_i$ = the HSPF predicted response which corresponded to the ith measured response sample.

- The variation among the model predictions not accounted for by the HSPF model would be computed as:

$$\sum_{i=1}^{I} \left(R_i - \hat{R}_i \right)^2$$

where R_i = the measured response value for the ith sample.

A standard R^2 value could then be computed to assess the level of fit of the HSPF model based on the monitoring data. The R^2 value would represent the proportion of the total variation among the observed data that could be explained by the HSPF model. For example, a relatively good fit of the HSPF model for freshwater discharge might be indicated by a high R^2 value of 0.75 which would indicate that 75% of the variation in freshwater discharge could be explained by the model.

In practice, R^2 values could be quickly computed using a statistical programming package such as SAS (SAS Institute, 1995) by computing one minus the ratio of the uncorrected sum of squares for the residuals over the corrected sum of squares for the measured values. Example SAS code to compute a measure of fit for freshwater discharge by primary basin would read as follows (SAS code in bold print):

- Step 1: Compile a SAS data set of primary-basin-specific HSPF flow estimates and the corresponding measured flows from the monitoring program, and compute residuals.

```
Data  set_a;
          input  basin  flowHSPF  flow;
          rflow = flow - flowHSPF;

         label   basin          = "Primary Basin"
                 flowHSPF       = "HSPF Predicted Flow (m3/sec)"
                 flow           = "Measured Flow (m3/sec)"
                 rflow          = "Residual of Predicted Flow (m3/sec)";
run;
```

- Step 2: Compute the corrected and uncorrected sums of squares for each basin.

```
Proc sort data=set_a; by basin; run;
```

```
Proc means data=set_a css uss noprint;
    var rflow flow ;
    output out=set_a css=css_r css_y uss=uss_r uss_y;
    by basin;
 run;
```

- Step 3: Compute and print a table of the R^2 values for each basin.

```
Data set_a;
        set set_a;
        Rsquare = 1 - (uss_r/css_y);
        label Rsquare = "R Square";
run;

Title1 "HSPF Model Run No. 123";
Title2 "Assessment of Model Fit";
 Proc print noobs label data=set_a;
        Var basin Rsquare;
run;
```

8.1.2. HSPF Model Validation

In addition to calibration of the HSPF model, the monitoring data could provide an important resource for validation of the model fit by setting aside a portion of the measured data before calibrating. These data would then be compared to the calibrated model responses to provide an independent measure of the fit of the model. In fitting any model to measured data, the fit of the predicted values to the observed values is incrementally improved by the inclusion of each additional explanatory parameter. In an extreme example, a model having the same number of parameters as points to be predicted will provide an exact prediction of the input data. However, each explanatory parameter added to a model decreases the robustness of the model to prediction error, and the model becomes more susceptible to producing biased results due to errors in estimation of the parameters. A complex watershed model such as the HSPF model is well saturated with independent explanatory parameters, and applying the same data used to calibrate the model as the only measure of its goodness of fit may not reveal prediction biases due to errors in parameter estimation.

An independent measure of the fit of the model could be provided by using a portion of the measured data from the monitoring program for validation of the HSPF model. These data would not be used to calibrate the model. The validation process would use the same combination of subjective and objective assessments of model fit as discussed for the calibration process, and would differ only in the measured data used.

Under certain conditions it may be desirable to combine the calibration and validation assessments into a single analysis. For example if the monitoring data collected for a particular basin would be too sparse to allow any large portion to be set aside for validation, then a standard “Jacknife” style statistical approach (Sokal and Rohlf, 1981) could be used for that basin. Using this approach, one or more observations would be set aside, the model would be calibrated using the remaining data, and the model’s predicted responses would be recorded. This process would be repeated by setting aside different groups of data until all of the available data had been used. The variation in parameter estimates and predicted values could then be assessed from the results of the repeated analyses. The disadvantages of applying the monitoring data to the HSPF model in this manner would be that the repeated calibration of the HSPF model would likely be an expensive and time consuming task, and the “Jacknife” method is very sensitive to outliers in the response data set such as would likely result from episodic storm events.

8.2. Land Use/Management Practices-Loading Relationships

One of the goals of the recommended short term intensive studies is to develop land use specific loading estimates for priority land use categories. These short term studies (one to five years) will serve to provide land use specific concentrations (and loadings) for defined basins in order to fill data gaps needed for improved HSPF modeling predictions and PLRG development. For the Estero Bay Watershed, we have prioritized land use categories for short term monitoring (Section 2.2) and have identified several specific secondary and tertiary basins containing those prioritized land use types (e.g., cropland, improved pasture, single family residential, etc.). It is particularly important that study basins have a uniform land use type, defined management practices, and a basin discharge point amenable to periodic flow measurements and water quality sampling.

The assessment of constituent concentrations (e.g., total suspended solids (TSS), total nitrogen, total phosphorus, etc.) from defined basins usually results in a high variation level, thus the constituents must be estimated based on probability (Horner, Skupien, Livingston, and Shaver, 1994). Therefore, to estimate an Event Mean Concentration (EMC) from a study basin, a large data set is needed to establish the underlying probability distribution for the locale or an assumption of the distribution and a smaller locale specific data set to fit the distribution.

The collection of empirical data will involve water quality sampling from a number of defined storms such that a typical annual cycle of storm types (of rainfall amount and duration) are represented in the sampled storms. Water quality samples should be analyzed for constituent concentrations measured periodically throughout the storm hydrograph such that representative concentrations are measured for a variety of flows. In order to calculate an Event Mean Concentration (EMC) for each storm event characterizing runoff from an identified basin, the concentration data must be flow-weighted. EMC can be calculated using the following equation:

$$EMC = \frac{\sum_{t=1}^{T} (C_t Q_t)}{\sum_{t=1}^{T} (Q_t)}$$

where T = total number of samples;

t = the time sample concentration © and flow (Q) are measured;

C_t = concentration measured at time t; and

Q_t = flow measured at time t.

Multiyear studies are preferable so year-to-year variability is taken into account. Upon completion of determining EMCs from a "representative" number of storms, mean concentrations can then be calculated for the constituents of interest for the defined land uses (and management practices) sampled. In the case of the Estero Bay Watershed, many of the flows are controlled and vary depending on specific water use needs within a basin. Therefore, in order to estimate loadings accurately, the acquisition of detailed records of the operation/water flow history from a basin may be necessary.

Land use specific loadings for a basin with mixed land uses can also be determined, if so desired, through use of the equations discussed and presented in the next section on rainfall-loading relationships for scenario 2. Use of these equations, however, will not account for land use specific management practices in the Estero Bay Watershed.

The impact of management practices on loadings from identified study basins can be predicted given the attenuation efficiencies of BMPs. The following describes two examples of how management practices (BMPs) may affect land use specific loadings to a water body. One example would be where loads generated and released from two basins, one downstream of the other, are each intercepted by a single BMP with known attenuation efficiencies. The final load delivered to the canal can thus be predicted by sequentially summing and subtracting in-stream loads given basin generated loads and BMP attenuation efficiencies. Another example would be where one basin delivers its load to two BMPs located downstream in series, which in turn, deliver a reduced load to the canal from the cumulative effect of attenuation.

The data measured by the monitoring program and the EMC samples could allow the statistical significance of the affects of land use or management practices on freshwater discharge and pollutant loading to be tested independently of the HSPF model. An analysis of covariance (ANCOVA) would be particularly well suited for this task. The question to be addressed by the analysis would be: can differences between freshwater discharge and pollutant concentrations between the sampled tertiary subbasins be attributed to land use or management practices. In order to address this question, one must account for the other important differences between the tertiary basins which could have affected discharge and loadings. The null hypothesis for this type of analysis would be that there is no significant difference between the response variable (e.g., freshwater discharge) which can be explained by differences in the main effect (e.g., proportion of tertiary basin in citrus land use) after taking into account other covariates (e.g. proportion of wetland coverage, rainfall).

The ANCOVA could be completed using the General Linear Models subroutines (GLM) of the SAS software package (SAS Institute, 1995). Using this software, the main effects (e.g., tertiary basins having different land uses to be compared), the response variable (e.g., TSS concentration), and the covariates (e.g., rainfall, proportion of secondary land use coverage such as wetlands) could be analyzed. An example ANCOVA design in SAS code would read:

```
Proc GLM;
   class basin;
   model TSS_conc = basin rainfall wetlands / solution;
   lsmeans basin / stderr;
run;
```

where

the variable **basin** holds values of each tertiary basin of relatively homogeneous land use to be compared,

the variable **TSS_conc** holds values of TSS concentration samples collected from the tertiary basins,

the variable **rainfall** hold values of rainfall recorded at the time of the TSS sampling, and

the variable **wetlands** holds values of the proportion of secondary wetland land cover in each tertiary basin.

The **stderr** option in the **lsmeans** line would cause the standard error of the least-squares means to be printed. The statistical probability that there is no difference between the response means would also be printed from this option.

In cases where the statistical tests for differences between measured values for two treatments (e.g., EMCs measured for two similar land use types) fail to reject the null hypothesis of no significant difference, the first conclusion would be that there was no difference observed between the two values. The next important question to be addressed is: how large of a difference could have existed between the two treatments and not have been detected by the monitoring program. Important differences between treatments could be missed by the monitoring program if any of the following are true:

- natural variability within the treatments was high relative to the magnitude of the difference between treatments,
- measurement variability was high relative to the magnitude of the difference between treatments, or
- the sample size of the monitoring program was not large enough given the variability within and among treatments.

The probability of a statistical test to detect a significant difference when a real difference exists between the treatments is the “Power” of the test. For difference tests such as the Student’s T tests, it would be necessary to determine the power of the tests for each test which failed to find a significant difference between treatments. One would then choose an acceptable probability of not detecting a real difference. A 10 or 5% probability of not detecting a real change would be reasonable for most pollutant loading estimates. The final step would be to compute the largest difference between the two treatments which could have been detected at the chosen probability level using standard power analysis methods for the test being used (Snedecor and Cochran, 1989). The size of this difference could then be judged as to whether it would be important or not. For example, a comparison of TSS concentration EMC data recorded for improved and unimproved pasture could result in a measured 20% difference in the concentrations between the two land uses. A T test of these results could indicate that this was not a significant difference. One would then conduct a power analysis to determine how large of a difference could have been detected given the sample size and levels of variability. The power analysis could indicate that only a 60% difference in the concentrations could have been detected given the variability in the EMC data and the sample size of the monitoring program. If the size of the smallest detectable difference is too large for management requirements, then the power analysis data could be used to determine the number of additional samples required to increase the power of the test to an acceptable level.

8.3. Rainfall-loading Relationships

Three possible scenarios in analyzing rainfall-loading relationships for the Estero Bay Watershed are discussed below. The selection of an approach for a basin will depend on the level of available data. The three conditions possible for available data are:

1) rainfall data, but no stream or canal flow and no water quality concentration data;

2) rainfall data, flow data, but no concentration data; and
3) rainfall data, and flow and constituent concentration data collected simultaneously.

The worst case scenario for this study area is that only rainfall data are available for an identified basin in which rainfall-loading relationships are needed. A better scenario is if both rainfall data and flow data are available. Finally, the best case for developing a rainfall-loading relationship is if rainfall, flow, and water quality or constituent concentration data are available. Each approach for the determination of rainfall-loading relationships is demonstrated through the following descriptions.

8.3.1. Scenario 1 (rainfall data, but no stream flow or water quality concentration data)

In this scenario, measured rainfall data can be used with land use specific runoff coefficients and runoff concentration data derived from the literature or from modeled rainfall-flow relationships from other watersheds of similar land use composition. Three approaches that can be used in the Estero Bay Watershed to develop rainfall-loading relationships when flow and concentrations data are lacking are discussed below.

The simplest approach entails the assumption of a linear relationship of rainfall to runoff. Many existing spreadsheet models have been developed to model stormwater runoff loadings using literature values for runoff coefficients and water quality concentrations. In these models, land use-specific runoff coefficients are used to estimate stormwater runoff in amounts proportional to the size of the basin, its land use and soils composition, and rainfall amounts. In many cases this method would be unacceptable because the relationship between rainfall and flow would be non-linear. A likely bias in the flow estimates is that low flows are overestimated and high flows are underestimated.

A non-linear regression can be developed to relate flow to rainfall. In this case, empirical data for flows must be available from a series of gages that contain similar land use composition as the basin of interest. A non-linear regression can then be developed between rainfall and known flow. This relationship can be developed to include monthly and seasonal parameters. Model complexity can be increased as needed until the desired level of confidence in the regression is reached. An acceptable relationship between rainfall and flow from the gaged basin can then be applied with rainfall in the basin of interest to estimate flows under the assumption that runoff characteristics are similar to the gaged basin used to develop the relationship.

A third approach is that flows and selected constituent loadings can be modeled from selected basins. Assuming detailed rainfall and watershed characteristic data are available for the basin, a more complex mechanistic model, such as HSPF or SWMM, could simulate flows and loads. This type of modeling approach has similar assumptions as the two previous approaches, but the algorithms used to develop flow and loadings estimates are much more complex.

8.3.2. Scenario 2 (rainfall and flow data, but no concentration data)

Given only rainfall and flow data for each time step of interest, basin specific loadings can be determined by apportioning the flow among the constituent land use categories and then estimating the load associated with each flow. To estimate land use specific pollutant loads, the fraction of the total subbasin flow that originated from each land use must be estimated. Assuming a monthly time step, the total subbasin flow can be apportioned among the constituent land use categories within each subbasin as follows:

$$\text{FLOW}_i = \frac{\text{FLOW} \times A_i \times R_i}{\sum_i A_i \times R_i}$$

where

FLOW_i	the total nonpoint source flow (cubic meters per month) from land use category *I*,
FLOW	the total nonpoint source flow (cubic meters per month) from a subbasin,
A_i	area (acres) in land use category *I*, and
R_i	a measured or literature based runoff coefficient (fraction of rainfall that runs off) for land use category *I*.

The above calculation would result in the disaggregation of the total nonpoint source flow for a basin into flows generated by individual land use types.

Basin pollutant loads can then be estimated by land use using land use/soil-specific runoff coefficients and land use-specific pollutant concentrations. These coefficients and concentrations can be derived from literature values or, as necessary, determined through the recommended short term land use specific monitoring programs in the Estero Bay Watershed previously described. Flows or runoff generated by an individual land use type within a basin can be calculated by prorating the total flow using land use-specific runoff coefficients and the area of that land use type within the basin as shown in the above equation. Loads for specific land uses can then be estimated for the desired time period according to the following equation:

$$\text{loading for land use } i = Flow_{LU_i} \times PC_{LU_i}$$

where:

PC_{LU_i} Average constituent concentration for land use i

$Flow_{LU_i}$ Total flow land use i

The land use-specific loads can then be summed to yield a total load for a basin. Basins within a larger area (secondary or primary basin) can be summed as appropriate to obtain the loading estimates at their desired spatial scale.

8.3.3. Scenario 3 (rainfall, flow and concentration data)

The best case scenario would occur when rainfall, flow, and concentration data would be available from the basin of interest. Given these available empirical data, or rainfall data and estimates of flow and loads, a rainfall-loading relationship can be completed to test the hypothesis: is loading from a specific basin related to rainfall during the current month or during some combination of previous months' rainfall. Regression analysis of rainfall amounts (independent variable) versus loads (dependent variable) can be used to mathematically describe rainfall-load relationships. Flows multiplied by concentration data for each time step will provide loading estimates.

Since the effects of rainfall measured in loadings downstream may be subject to varying time lags depending on basin size, land use, and previous rainfall amounts, iterative regressions using one or more rainfall variables to determine the "best fit" relationship is recommended. For example, rainfall measured one month, two months, or one and two months (as separate independent variables) prior to the time of loadings measurements could be iteratively regressed to find the best fit relationship. The prediction of basin specific loads for combinations of months (assuming a monthly time step) and basins can be calculated as:

$$Load_{bst} \qquad \alpha_{b,s} \qquad \beta_b X_{bst} \qquad \gamma_b X_{bst\ 1}$$

where

Load$_{bst}$ = predicted load for basin *b* for season *s* using time step *t*;

$\alpha_{b,i}$ = y intercept for basin *b* and season *s*;

β_b = slope of regression for basin *b* using current month's rain;

γ_b = slope of regression for basin b using pervious month's rain;

X$_{bst}$ = current month's rain; and

X_{bst-1} = previous month's rain.

SAS code (in bold) applicable to calculating basin loads using the above equation is shown below.

```
Proc glm;
   class month basin;
   model load = month season rain(basin)
          lagrain(basin) / solution;
run;
```

8.4. Other Analyses

The data collected from the recommended monitoring programs are applicable to other analyses that may be useful in developing PLRGs for the Estero Bay Watershed. These analyses include the assessment of spatial and temporal trends and the determination of relationships between stressor and response variables.

8.4.1. Spatial and Temporal Analyses

It is anticipated that spatial and temporal data analysis will be applicable to rainfall, flow, concentration, and loading data collected at the identified long term monitoring program sites and at the 25 sampling sites initially suggested for synoptic sampling. The following discussion will include analyses of spatial and temporal relationships that could be applied to the monitoring data.

8.4.1.1. Approach to Assessing Spatial and Temporal Trends

The assessment of spatial and temporal trends in monitored parameters (e.g., rainfall and loading parameters) will provide important information for the PLRG development process. Such analyses could also facilitate the selection of management targets appropriate for current and expected future conditions.

The analysis of spatial and temporal trends in collected data will necessitate the computation of means and variances for the various levels of space and time of concern, such as regional, annual, seasonal, monthly, or daily levels. Long-term continuous data records, such as gaging station data, could be analyzed for trends at time scales reflective of the collection frequencies for the entire period of record (daily, weekly, monthly, etc.). The possible discontinuous nature of historical water quality data, or data collected from a synoptic study, for example, could lead to an analysis with greater emphasis on the spatial relationships between similar parameters and/or among different parameters. Analysis of data sets covering different time periods could indicate long-term changes in parameter concentrations for localized areas. Such analysis will illustrate the degree that water quality conditions have changed over a relatively long period of time.

Other potential approaches include the plotting of mixing curves of selected parameter concentrations over a gradient of concentrations along the longitudinal axis of a canal reach. Spatial patterns of variability can be summarized through use of contour mapping techniques showing isopleths of summary statistical values (e.g., annual mean, climatological mean, and variance in parameter concentrations for a given location).

Analyses of existing data could also assess spatial and temporal trends in nutrient and pollutant loadings and ambient water quality parameters in a statistically rigorous manner. For example, methods based on the general linear model of classical statistics (Neter et al., 1985), such as analysis of variance, regression, and analysis of covariance could be used to detect spatial and temporal trends in univariate water quality response parameters. Analysis of variance techniques could be used to detect regional or year to year differences in water quality. Regression techniques could be used to test for a particular type of trend, such as a linear trend over time in water quality and/or nutrient loading. Analysis of covariance could be of use in detecting trends, while removing the effect of a secondary source of variability in the response. For example, analysis of covariance could be used to detect annual changes in total nitrogen concentration that would have existed if freshwater flow in all years was constant. Differences in total nitrogen could then be attributed to sources other than flow.

Box-Jenkins time series analysis (Box and Jenkins, 1976) could be applied to assess short term or long term temporal trends at a particular site. In these techniques, models could be fit to a time series of observations that are equally spaced in time (e.g., monthly averages), and tests of significant seasonal or long term trends could be undertaken. Transfer functions could also be used in Box-Jenkins analysis to include the effects of other parameters on the response of interest.

Classical statistics and Box-Jenkins analysis necessitate the assumption of normally distributed data for most applications. Nonparametric or distribution-free methods (Hollander and Wolfe, 1973) may be used when the assumption of normally distributed data is not met. These methods generally involve the use of a rank transformation of the data, and the application of classical techniques to the ranks. Examples of nonparametric methods include Spearman's rank correlation and Kendall's tau for assessing temporal trends, and Wilcoxon's rank analysis of variance and Friedman's two-way analysis of variance for assessing year-to-year and spatial differences in water quality.

Most of the statistical techniques mentioned thus far require observations to be independent of each other. In estuarine and freshwater monitoring programs, observations are often correlated with each other. For example, a total nitrogen concentration obtained at a specific location and time may be correlated with the total nitrogen concentration at the same location one month earlier (temporal correlation), or with total nitrogen concentrations at nearby locations within the canal reach (spatial correlation). Methods which can be used to analyze spatially correlated data include variogram estimation and krieging (Cressie, 1991). Statistical analyses based on the assumption of independent data when applied to correlated data may result in erroneous results. For example, statistical analyses used to detect long term trends in total nitrogen which do not account for short term temporal and spatial correlation in the total nitrogen concentrations may lack the minimal power needed to detect such trends. Therefore, statistical methods that are applied to environmental monitoring data should account for spatial and temporal correlation in the observed data.

If all of the necessary parameters can not be successfully estimated based on historical or existing monitoring data, trend assessments could be undertaken based on nonparametric tests of trend. For example, Kendall's tau could be used to assess temporal trends in total phosphorus concentrations at a specific site, or for the average condition across a group of sites. Estimated long term trends based on Kendall's tau can be combined from different seasons or sites to estimate an overall trend in nitrogen concentrations (Hirsch et al., 1982, Hirsch and Slack, 1984). Tests could be applied to Kendall's tau estimates from different seasons or sites to determine if a trend is homogeneous, or if different seasons or sites within a river or estuary are experiencing different trends (Van Belle and Hughes, 1984).

A more detailed and specific approach to data analysis can be developed when the specific monitoring programs to be implemented are identified. Data analysis could emphasize the following elements:

- temporal (monthly, seasonal, annual, multi-year) relationships in surface and groundwater discharge, and rainfall;
- temporal and spatial trends in loadings of macronutrients (e.g, nitrogen and phosphorus forms) and other significant water quality constituents (e.g., biological oxygen demand, total suspended solids) to the Estero Bay estuary or other identified waterbodies or canal reaches of concern; and
- temporal and spatial trends in concentrations of nutrients, chlorophyll, dissolved oxygen, biological oxygen demand, suspended solids, and other water quality indicators (e.g., Secchi disk disappearance depth, turbidity, light attenuation) within the identified waterbodies.

8.4.2. Analyses of Long Term Trends

Methods for characterizing long term empirical trends in the monitoring data could include the plotting of long-term data records and testing for trends through statistical analyses. The examples of statistical analyses described below were based on the application of the general linear model using least squares estimation methods.

Rainfall

At selected representative precipitation stations, time series plots of mean monthly precipitation for each year could be constructed for each calendar month of the year (i.e., plot of mean rainfall in January for each year for the period of record). Time series of data at specific locations could also be examined to assess differences in rainfall within the Estero Bay Watershed. The objective of these analyses would be to characterize long-term trends in rainfall within the watershed or basin of interest.

Rainfall data from representative monitoring stations in the Estero Bay Watershed should be identified for the specific basin or basins of interest. Trend analyses can then be conducted separately by basin and season using the following linear model:

$$Y_{ij} = \mu + \alpha_{1,i} + \alpha_{2,j} + \epsilon_{ij}$$

where

Y_{ij} = total seasonal rainfall at station j in year I

μ = overall mean basin-wide seasonal rainfall

$\alpha_{1,i}$ = deviation from the mean rainfall in year I

$\alpha_{2,j}$ = deviation from mean rainfall for station j

ϵ_{ij} = random error term.

The annual effects on rainfall ($\alpha_{1,i}$) can be estimated using least squares methods, and a linear contrast of these estimates could be calculated to assess linear trend in rainfall. A T test based on this contrast could be performed to assess the significance of the trend in rainfall.

Flows

The analysis of long term trends in flow is similar to rainfall. Time series plots of periodic flows (daily, weekly, monthly, etc.) and average discharges for representative gage sites should be produced and examined. Following the completion of a set of representative time series plots, specific statistical tests could be applied to selected gaged flows.

For example, statistical analyses could be conducted to assess year to year changes in flow at stations in the Estero Bay Watershed. Trend analyses for canal discharge could be conducted separately by station and season using the following linear model:

$$Y_i \quad \beta_0 \;\; \beta_1 i \quad \epsilon_i$$

where

Y_i = total seasonal discharge (meters/season) in year I

β_0 = intercept term

β_i = annual trend for seasonal discharge

ϵ_i = random error term.

In this linear model, flows can be converted to a volume of water for an identified season, divided by the drainage area that contributes flow to the gaging station of interest. The resultant units will be in terms of the height of water (for the drainage area) per unit time (season).

Estimates of the annual trend in canal discharge (β_1) can be calculated based on least squares estimation, and the significance of the estimated trend (e.g., hypothesis = There is a significant increasing trend in canal discharge over time.) can be assessed using a T test statistic.

Water Quality and Loads

An analysis to assess long term spatial and temporal trends in water quality and loadings could provide important information for the PLRG development. Flow and water quality data, for example, can be performed using a monthly time step. Flow data could be combined with water quality data to determine loads of selected constituents. Flow data can be calculated as mean daily estimates of canal or stream discharge in cubic feet per second (cfs). Monthly arithmetic averages (cfs) can also be calculated from the daily flow estimates, and those monthly values can be used in all subsequent data analyses.

A series of steps can be performed to select the stations and water quality variables used for the trend analyses. Initially, a list of all variables and their overall frequency in the water quality data set should be constructed. From this list, a subset of variables (at a minimum TSS, total nitrogen, total phosphorus) can be selected.

All canal stations within the study area could then be examined for their temporal coverage of the selected water quality variables with the objective being to use these data to assess long-term spatial and temporal trends in water quality and loads. In order to eliminate sites with too few records to adequately assess long term trends, sites with fewer than a certain number of measurements overall (e.g., 10) should be removed from consideration. Subsequent examination should include a listing of minimum and maximum dates of sampling, as well as the average interval between sampling events to ensure consistency among sites compared. The temporal record of flow measurements should also be examined. Stations with data for a sufficiently long period of record for both the selected water quality variables and flow should be retained for further screening.

Stations can be further subset for analysis by examining plots of water quality measurements over time. Stations with large gaps of missing data between the first and last dates sampled should be removed from the analyses. A rule should be established with regard to the minimum number of data points that are representative of an annual cycle. Temporal plots (i.e., plots of water quality measurements versus date) can also be used, along with histograms or “box and whisker” plots, to characterize data distributions and to identify outliers. Obvious outliers (e.g., temperatures above 40^{o}C; salinities greater than 40 ppt) should be removed before further analyses are performed.

For the identified gages with flow and water quality data, mean (period of record) wet and dry season water quality constituent concentrations and loads can be compared graphically. Time series plots of concentrations and loads could also be produced for each gage site. The plots of empirical data should be carefully examined prior to performing statistical tests for long-term trends. The visual examination of time series plots should provide an indication of which sites have definable trends with time for the selected parameters.

Assessments of annual trends in canal or tributary loads and water quality conditions can be conducted separately by site and season using the following model:

$$Y_{id} \quad \mu \quad "_{1,i} \quad "_{2,d} \quad ,_{ij}$$

where

Y_{id} = reported concentration (e.g. mg/l) or calculated load (tons/month) for year I and month d

μ = overall mean concentration or load

$"_{1,i}$ = deviation from the mean in year I

$"_{2,d}$ = deviation from mean for month d

$,_{id}$ = random error term.

The annual effects on concentrations or loads ("$_{1,i}$) can be estimated using least squares methods, and a linear contrast of these estimates can be calculated to assess linear trend in concentration or load. The within year effects on loads or concentrations ("$_{2,d}$) can be included in the model to remove within season variation from the residual error, and thus improve the power of testing for annual trends. A T test based on the linear contrast of annual effects can be conducted to assess the significance of the trends in load and concentrations. For example, the hypothesis -- there is a significant increasing trend in total nitrogen loading with time -- could be tested.

8.4.3. Identification of Relationships Between Stressor and Response Variables

Correlation statistics can be used to measure the association between two continuous variables. Linear correlation coefficients (comparison of independent variables) can be used to test for potential relationships between stressor variables (nutrient or pollutant loading variables) and response variables (water quality parameters) in the receiving waterbodies (canal reaches or estuarine areas) of those loads. Once data are formatted into SAS data sets, correlation coefficients between loading variables and suspected response variables can quickly and efficiently be calculated. For example, total nitrogen loads from the Ten-Mile Canal Basin could be correlated with ambient water quality response variables (e.g., chlorophyll-a, trophic state index, light attenuation) measured in Estero Bay that may be expected to change (increase or decrease) as the result of changing loads in nitrogen. Examples include Pearson product-moment correlation coefficients and Spearman coefficients (SAS, 1988). Data used for correlation statistics require certain assumptions about the paired variables (e.g., X and Y). In correlation, the Y's at each X, and the X's at each Y, are assumed to come at random from a normal distribution. This is referred to as sampling from a "bivariate normal distribution." Substantial correlation among variables can be adversely affected by nonnormality, thus data subjected to correlation statistics must be checked for normality prior to testing.

In addition to correlation statistics, regression models can be used to statistically test for relationships between the variation in a response variable to the variation in one or more stressor variables (loadings). For example, a model can be developed to relate total nitrogen loads to the Imperial River to the response variables of chlorophyll-*a* and turbidity measured in identified segments of the Estero Bay. Regression techniques can be used to test both linear or non-linear relationships. A significant statistical relationship among such variables could provide a management tool which could be used to predict phytoplankton biomass, measured as chlorophyll-*a* concentration, for example, based on total nitrogen load to the waterbody. If a target chlorophyll-*a* concentration can be set for the receiving waterbody based on living resource requirements, then PLRGs for nitrogen could be determined using such a modeling approach.